

XIV Encontro Nacional de Pesquisa em Ciência da Informação - XIV ENANCIB 2013

GT 8: Informação e Tecnologia

Modalidade de apresentação: Comunicação Oral

O PAPEL DAS ONTOLOGIAS NA INTEROPERABILIDADE DE SISTEMAS DE INFORMAÇÃO: REFLEXÕES NA ESFERA GOVERNAMENTAL

Fernanda Farinelli
PPGCI-ECI/UFMG

Stefane Melo
PPGCI-ECI/UFMG

Maurício B. Almeida
PPGCI-ECI/UFMG

Resumo

A questão da interoperabilidade entre sistemas de informação está já há muitos anos na agenda de diversos governos ao redor do mundo, incluindo o Brasil. Prover serviços governamentais de qualidade requer da adoção de soluções visando interoperabilidade, de forma a possibilitar a integração de sistemas e o compartilhamento de informações entre as várias instâncias de governo. Em alguns setores governamentais, como por exemplo, de atenção à saúde, a interoperabilidade é uma necessidade. O presente artigo é uma iniciativa no sentido de discutir essa questão no âmbito da Ciência da Informação. Em particular, pretende-se esclarecer o papel as ontologias nesse contexto, uma vez que ontologias vem sendo recomendadas como alternativa para prover interoperabilidade. Para tal, define-se interoperabilidade, apresentam-se iniciativas do governo brasileiro nesse sentido, para finalmente elucidar a aplicação de ontologias em interoperabilidade. Apresentam-se ainda resultados parciais de pesquisa relacionada em andamento envolvendo ontologias e modelagem de sistemas. Conclui-se que ontologias podem desempenhar papel relevante na busca pela interoperabilidade entre sistemas, mas que muito ainda está por ser feito para que essa possibilidade se torne real.

Abstract

Interoperability among information systems is an issue that has been present in the agenda of several governments around the world, including the Brazilian government, since many years ago. In order to provide high quality public services, the public sector needs to pursue solutions towards interoperability, as a way to enable the integration of information systems and the sharing of information across the several governmental levels. Within some public sectors, as for example those ones concerned with basic healthcare, interoperability is a real requirement. This article is an initiative for discussing the issue of interoperability in the scope of Information Science. Particularly, we intend to clarify the role of ontologies in this context, insofar as ontologies have been proposed as an alternative to provide interoperability. In order to reach our goals, we briefly define interoperability, present the initiatives of the Brazilian government in this sense, and then try to elucidate how to apply ontologies as a tool to provide interoperability. In addition, we present partial results of ongoing research involving ontologies and the improvement of systems modeling. We conclude that ontologies can perform a significant role in mitigating problems of interoperability, but considering the current stage of Brazilian initiatives, there are many things to be done before we can reach real results.

1) Introdução

A heterogeneidade dos dados e informações na Administração Pública desafia a cooperação e a sistematização de processos no escopo federal, estadual e municipal. A massiva disseminação da informação alavancada pela crescente disponibilidade de tecnologias de informação, como as redes, os computadores, a web, e os sistemas de informação (SIs), tem gerado dificuldades cada vez maiores para integração das fontes de informação das instâncias governamentais.

Neste sentido, as dificuldades de busca, acesso, armazenamento e recuperação da informação vêm se tornando mais e mais complexas. O grande desafio enfrentado pelo governo nas diferentes esferas é a integração de diferentes tipos de informação, tanto com relação ao conteúdo quanto a natureza. De fato, a falta de padronização consistente impede a interoperabilidade entre SIs governamentais.

Prover serviços governamentais eficientes e de qualidade requer a adoção de soluções visando interoperabilidade, de forma a possibilitar a integração dos sistemas e o compartilhamento das informações entre os vários órgãos e instâncias de governo.

Vários países, como Estados Unidos, Canadá, Reino Unido, Austrália e Nova Zelândia, vêm adotando medidas para modernização da administração pública por meio da integração de serviços públicos. O governo brasileiro tem acompanhado esta tendência internacional, ao buscar definir padrões para mitigar problemas de integração entre SIs e promover interoperabilidade. Em alguns setores governamentais, como por exemplo, o de atenção à saúde, a interoperabilidade é mais do que um desejo, é uma necessidade.

O presente artigo é uma iniciativa no sentido de discutir essa questão no âmbito da Ciência da Informação. Em particular, pretende-se esclarecer que papel que as ontologias desempenham nesse contexto, uma vez que esse instrumento é recomendado como alternativa para promover interoperabilidade (SMITH, 2004; GUARINO, 1998). Busca-se então responder a questão: de que forma ontologias podem proporcionar interoperabilidade?

Para tal, inicialmente, revisitam-se aspectos relevantes para o entendimento da noção de interoperabilidade e da importância que a questão tem assumido (seção 2). Em seguida, discorre-se brevemente sobre iniciativas nesse sentido no âmbito do governo brasileiro (seção 3). Cabe destacar que não se pretende uma revisão exaustiva das iniciativas nacionais e internacionais, mas apenas o necessário para caracterizar o atual

estado da arte nacional. Em seguida, busca-se elucidar a aplicação de ontologias em interoperabilidade, além de apresentar um estudo de caso de pesquisa em andamento (seção 4). Finalmente, conclui-se que ontologias podem desempenhar papel relevante na busca pela interoperabilidade entre sistemas, mas que muito ainda está por ser feito para que essa possibilidade se torne real (seção 5).

2) Interoperabilidade.

A busca por interoperabilidade na área de tecnologia da informação não remete a algo novo, mas tem se tornado uma necessidade à medida que a web derrubou as barreiras organizacionais, conectando bancos de dados corporativos antes isolados.

O *Merriam-Webster Dictionary* define interoperabilidade como a habilidade de um sistema em utilizar partes de outro sistema. O *Institute of Electrical and Electronics Engineers* (IEEE)¹ sugere que interoperabilidade é a capacidade de um sistema, ou de um produto, para funcionar com outros sistemas, ou produtos, sem nenhum tipo de esforço especial por parte de um cliente. Para o IEEE a interoperabilidade é possível apenas a partir da aplicação de normas e do uso de padrões. Nesse sentido, a interoperabilidade está atrelada a cooperação, normalizada por especificações, políticas e padrões que viabilizem o intercâmbio integrado de informações.

Na literatura de Ciência da Informação e de Ciência da Computação encontram-se exemplos de referências que definem interoperabilidade, de forma simples, como a capacidade que um sistema possui de compartilhar e trocar informações e aplicações com outro sistema (BISHR, 1997; SHETH, 1999; SAYÃO e MARCONDES, 2008).

Marcondes e Sayão (2001) explicam interoperabilidade como a possibilidade do usuário buscar por recursos informacionais heterogêneos, armazenados em diferentes locais de uma rede, utilizando-se de uma interface única e sem necessidade de conhecimento sobre como os recursos estão armazenados. Arms (2002) explica que a interoperabilidade tem como objetivo desenvolver serviços e soluções úteis para os usuários, a partir de recursos informacionais que são tecnicamente diversos, e muitas vezes gerenciados por instituições diferentes. Sheth (1999) sugere problemas de interoperabilidade podem estar relacionados a fontes de informação manipuladas pelos sistemas, as quais podem apresentar diferenças sintáticas, estruturais ou semânticas. Para Cruz (2005) citado por Santos (2011), a questão sintática diz respeito ao uso de

¹ Disponível na Internet em: <www.ieee.org/>. Acesso em: 14 de junho de 2013.

diferentes modelos ou linguagens; a questão estrutural está relacionada a divergências entre as estruturas de dados adotadas por cada sistema; a questão semântica remete a adoção de interpretações divergentes para a informação intercambiada entre os sistemas.

Para mitigar esses diferentes tipos de problemas, Arms et al. (2002) estabelecem que o grau de cooperação entre sistemas deve ser diferenciado em pelo menos três níveis de cooperação:

- *Acordo técnico*: busca promover interoperabilidade tecnológica, através da uniformidade da informação e dos serviços utilizados por dois ou mais sistemas; Envolve a utilização de formatos, protocolos, e padrões de forma que mensagens possam ser trocadas entre diferentes plataformas;
- *Acordo sobre conteúdo*: busca promover interoperabilidade semântica e para isso se valem de desenvolvimentos da representação e organização do conhecimento; envolve o uso de metadados e recursos para uniformizar a interpretação de mensagens;
- *Acordo organizacional*: busca reduzir diferenças políticas, através da reunião das organizações em acordos federativos com o intuito de implementar padrões e tecnologias comuns; envolve regras básicas para acesso, alteração e autenticação da informação, bem com integração entre serviços.

Quando se pretende que dois ou mais SIs colaborem, é preciso que possam interoperar. Isso envolve capacidade de comunicação, de troca de informações, de uso de operações mutuamente, de forma independente das arquiteturas, plataformas e semânticas utilizadas. Os problemas de integração também se revelam em outros níveis que extrapolam o SI e abrangem variáveis contextuais, tornando a questão ainda mais complexa (MILLER, 2000). Um cenário amplo, envolvendo outros níveis onde se deve buscar acordo, define tipos de interoperabilidade a se buscar (UKOLN, 2005):

- *Interoperabilidade técnica*: abrange padrões de comunicação, de transporte, de armazenamento e de representação de informações;
- *Interoperabilidade semântica*: refere-se ao significado da informação originada em diferentes sistemas; envolve a adoção de soluções capazes de assegurar interpretações uniformes entre os sistemas, como por exemplo: esquemas de metadados, classificação, tesouros e ontologias;
- *Interoperabilidade organizacional*: relacionada ao contexto organizacional, envolve fluxos de trabalho e de informação, as relações de poder e a cultura da instituição; através da modelagem de processos de negócio, busca alinhamento entre informações presentes na arquitetura corporativa;
- *Interoperabilidade política e humana*: envolve a forma com a informação é disseminada, e a decisão consciente de torná-la disponível na organização;

- *Interoperabilidade intercomunitária*: aborda o acesso a informações originadas em diferentes fontes, por organizações, especialistas e comunidades de natureza distintas; remete à interação entre domínios independentes;
- *Interoperabilidade legal*: relacionada a exigências e implicações legais de tornar a informação livre e amplamente disponível;
- *Interoperabilidade internacional*: envolve a cooperação em escala internacional, onde o intercâmbio envolve uma grande diversidade de padrões e normas, além de problemas inerentes de comunicação por barreiras linguísticas.

Lara e Moreira (2012) explicam que a interoperabilidade, vista como meio de promover cooperação, torna evidente a necessidade de definir políticas que vão viabilizar a produção e a recepção de diferentes tipos de informação. Como parte da definição dessas políticas de informação, Landsbergen e Wolken (2001) sugerem a criação e a adoção de normas e padrões visando à interoperabilidade, os quais vão proporcionar efetividade, eficiência e responsividade.

De fato, nos últimos anos tem se observado a proliferação de políticas, padrões e normas em países como Estados Unidos, Canadá, Reino Unido, Austrália e Nova Zelândia. A criação desses instrumentos segue uma tendência de modernização da administração pública para, em última instância, proporcionar melhores serviços ao cidadão. A integração de serviços públicos em todas as esferas de governos depende incondicionalmente da adoção de medidas que promovam a interoperabilidade dos serviços de governo eletrônico (SANTOS, 2010).

No Brasil, a proposta não é muito diferente. Busca-se interoperabilidade a partir da soma de vários esforços, como a integração de sistemas, a integração de redes, a troca uniforme de dados e a definição de tecnologia, levando-se ainda em consideração a existência diversas arquiteturas, plataformas de *hardware* e *software*.

3) Interoperabilidade e as iniciativas brasileiras.

Assim como ocorre em países mais industrializados, o governo brasileiro está envolvido em uma iniciativa de governo eletrônico desde 2004. O governo federal patrocina uma arquitetura de interoperabilidade de serviços para governo eletrônico denominada *e-PING - Padrões de Interoperabilidade de Governo Eletrônico*, a qual define um conjunto mínimo de premissas, políticas e especificações técnicas que regulamentam a utilização de tecnologias da informação e comunicação. Pretende-se assim promover interoperabilidade no âmbito do instituições públicas estabelecendo-se

condições de interação entre a sociedade e as instituições governamentais, sejam federais, estaduais ou municipais.

Segundo Santos (2010), a arquitetura e-PING teve como modelo o projeto *e-GIF - Government Interoperability Framework* conduzido pelo governo britânico desde o início dos anos 2000. Os padrões da e-PING envolvem cinco segmentos:

- **Interconexão:** estabelece condições para que órgãos de governo se conectem, além de fixar as condições de interoperação entre governo e sociedade;
- **Segurança:** trata dos aspectos de segurança para assegurar a validade e privacidade das operações;
- **Meios de acesso:** define as questões relativas aos padrões dos dispositivos de acesso aos serviços de governo eletrônico;
- **Organização e intercâmbio de informações:** aborda o gerenciamento e a transferência de informações nos serviços de governo eletrônico;
- **Integração com o Governo Eletrônico:** estabelece diretrizes para intercâmbio de informações baseados nas definições e-PING.

No âmbito do e-PING, o governo brasileiro promove outra iniciativa denominada *e-VoG - Vocabulários e Ontologias do Governo Eletrônico*. O e-VoG consiste de um conjunto de padrões, ferramentas e metodologias que tem por objetivo promover interoperabilidade semântica no e-PING (ePING, 2013). O e-VoG engloba uma taxonomia especificada no *Vocabulário Controlado do Governo Eletrônico (VCGE)* e o *Padrão de Metadados do Governo Eletrônico (e-PMG)*.

O VCGE é uma estrutura hierárquica continuamente incrementada, criada pelo Ministério do Planejamento, Orçamento e Gestão com o objetivo de facilitar a navegação das pessoas nos portais do Governo Federal. Pretende auxiliar usuários de sites governamentais a encontrar informações, mesmo considerando que não tenham conhecimento de qual órgão estaria envolvido (ePING, 2013; VCGE, 2011). A Figura 1 apresenta um fragmento do esquema da representação de termos do VCGE.

Figura 1: Níveis de representação dos termos do VCGE.



Fonte: VCGE (2011)

O VCGE está organizado em uma hierarquia que permite herança múltipla, de forma que certos conceitos pertençam simultaneamente a mais de um domínio. A Figura 2 apresenta um fragmento de termos no VCGE.

Figura 2: Termos representados no VCGE.

VCGE - Vocabulário Controlado do Governo Eletrônico	
A	
- Abastecimento TG: Agricultura, extrativismo e pesca TE: Armazenamento de alimento TE: Comercialização agrícola TE: Estocagem TE: Segurança alimentar TE: Suprimento de alimento TE: Transporte	TE: Convenção internacional TE: Tratado internacional
	- Acupuntura TG: Medicina natural
	- Adaptação à mudança do clima TG: Mudanças climáticas
	- Aditivos de alimentos TG: Tecnologia de alimentos
	- Administração de concessões TG: Administração financeira

Fonte: VCGE (2011). TG = termo genérico; TE = termo específico

O e-VoG define também o chamado *e-PMG - Padrão de Metadados do Governo Eletrônico*, o qual se propõem a definir a semântica dos elementos e dos qualificadores para descrição de recursos informacionais. Tal padrão constitui um conjunto mínimo de elementos que referenciam os dados necessários para a recuperação e gerenciamento de informações. O objetivo e-PMG proporcionar o acesso a descrições dos recursos para a pesquisa de informações do governo brasileiro na web (ePMG, 2010). A Figura 3 mostra um fragmento de um elemento descrito pelo e-PMG.

Figura 3: Exemplo de elemento e-PMG

3.9 Destinatário	
Nome	Destinatário
Identificador	<i>Addressee</i>
Definição	Entidade (pessoa ou organização) para quem a informação contida no documento foi dirigida.
Obrigatoriedade	Obrigatório se aplicável
Objetivo	Identificar a entidade (organização ou pessoa) a quem o recurso foi dirigido. Demonstrar a autenticidade de um documento, indicando a quem o documento é dirigido.

Fonte: ePMG (2010)

A portaria num. 05 de 14 de Julho de 2005 do governo federal torna obrigatória a adoção dos padrões e políticas descritos na arquitetura e-PING para todos os órgãos do governo federal. A arquitetura e-PING ainda inclui um *Guia de Gestão de Processos de Governo* que objetiva sugerir, do ponto de vista conceitual, as melhores práticas sobre processos no âmbito público (GPG, 2011).

Cabe ainda destacar o esforço do governo na área de saúde, onde existem iniciativas de padronização para SIs aplicados a saúde, como por exemplo, aquela institucionalizada pela portaria num. 2073 de 31 de agosto de 2011 do Ministério da

Saúde. Esse documento estabelece um conjunto de padrões para interoperabilidade da informação em saúde, abrangendo os SIs do SUS, em níveis municipal, distrital, estadual e federal. Além disso, institui a certificação para sistemas de registro eletrônico em saúde criada a partir de uma parceria técnico-científica entre a Sociedade Brasileira de Informática em Saúde e com o Conselho Federal de Medicina (BRASIL, 2011).

A portaria estabelece ainda que os padrões de interoperabilidade constem do *Catálogo de Padrões de Interoperabilidade de Informações de Sistemas de Saúde* (CPIISS), os quais são publicados pelo Departamento de Informática do SUS, e devem estar disponíveis para a sociedade. O CPIISS segue o formato definido pelo e-PMG, estabelecendo o uso das seguintes tecnologias:

- Para interoperabilidade entre sistemas SUS: *Web Services*, no padrão SOAP 1.1;
- Para segurança e integridade: *WS-Security* para criptografia e assinatura digital;
- Para identificação de *Web-Services*: *Uniform Resource Identifier* (URI);
- Para descrição de *Web-services*: *Web Service Description Language* (WSDL);
- Para Registro Eletrônico em Saúde (RES): o padrão OpenEHR²;
- Para os resultados e solicitações de exames médico: o padrão HL7³;
- Para codificação de termos clínico: SNOMED-CT⁴;
- Para a interoperabilidade com a saúde suplementar: padrão TISS⁵;
- Para a definição da arquitetura do documento clínico: padrão HL7 (CDA);
- Para a resultados e solicitações de exames de imagens: DICOM⁶;
- Para a codificação de exames laboratoriais: LOINC⁷;
- Para a codificação de dados de etiquetas de sangue: ISBT 128⁸;
- Para a interoperabilidade de modelos de conhecimento: ISO 13606-2;
- Para cruzar identificadores de pacientes: IHE-PIX.

4) Ontologias como alternativa para a interoperabilidade

Ontologias têm sido descritas como soluções para interoperabilidade. Entretanto, nem sempre está claro que tipo de solução é essa que as ontologias podem proporcionar. Essa seção se ocupa dessa questão. Apresentam-se inicialmente algumas características básicas das ontologias (seção 4.1), para em seguida descrever como ontologias podem ser úteis na solução de ambiguidades inerentes ao processo de troca de informação

² Disponível na Internet em: <<http://www.openehr.org/home.html>>. Acesso em: 29 de julho de 2013.

³ Disponível na Internet em: <<http://www.hl7.org/>>. Acesso em: 29 de julho de 2013.

⁴ Disponível na Internet em: <<http://www.ihtsdo.org/snomed-ct/>>. Acesso em: 29 de julho de 2013.

⁵ Disponível na Internet em: <<http://www.ans.gov.br/>>. Acesso em: 29 de julho de 2013.

⁶ Disponível na Internet em: <<http://medical.nema.org/>>. Acesso em: 29 de julho de 2013.

⁷ Disponível na Internet em: <<http://loinc.org/>>. Acesso em: 29 de julho de 2013.

⁸ Disponível na Internet em: <<http://www.iccbba.org/>>. Acesso em: 29 de julho de 2013.

(seção 4.2). Finalmente, apresenta-se um estudo de caso de pesquisa em andamento sobre o uso de ontologias na modelagem conceitual de sistemas (seção 4.3).

4.1) Ontologias e sua aplicação em interoperabilidade

Ontologias tem sido objeto de estudo em diferentes campos de pesquisa. É possível encontrar publicações sobre ontologias em Filosofia, Ciência da Computação e Ciência da Informação, aplicadas a domínios ainda mais diversos como medicina, biologia, engenharia, geografia e direito. Trata-se, portanto, de assunto interdisciplinar e para tirar proveito do que as ontologias têm a oferecer é preciso entender o sentido do termo nos diversos campos de pesquisa (ALMEIDA, 2013).

Em Filosofia, ontologia diz respeito aquilo que existe e tem sido estudada desde a antiguidade, inicialmente nos trabalhos de Aristóteles. O objetivo, em última instância, é entender o mundo, suas entidades e as relações entre essas entidades. Em Ciência da Computação, as ontologias são consideradas artefatos de engenharia de software com diversos usos (GRUBER, 1993). São utilizadas em modelagem de SIs, onde funcionam como um tipo de metamodelo; são usadas em Representação do Conhecimento, onde funcionam como uma representação legível por máquina para fins de inferência automática. Em Ciência da Informação, princípios ontológicos têm sido utilizados desde o século XIX, em estudos bibliográficos para representação do conteúdo de documentos (VICKERY, 1997). Nesse contexto, ontologias são um tipo de sistema de organização do conhecimento (KOS)⁹ que possibilita representação do conhecimento (SOERGEL, 1997).

No início dos anos 1990, ontologias passaram a ser amplamente aplicadas nas áreas de medicina e biomedicina como forma de estruturar o grande volume de dados gerados. Desde então, essas áreas têm abrigado pesquisa sobre interoperabilidade de SIs a partir de ontologias (Simon e Smith, 2004), como demonstram as inúmeras iniciativas internacionais produzidas com essa tecnologia¹⁰.

No Brasil, a preocupação com a interoperabilidade de sistemas médicos ficou evidente através da portaria 2.073 (Ministério da Saúde, 2011) já mencionada anteriormente (seção 3). Uma das recomendações dessa portaria é justamente adotar ontologias para lidar com a questão da interoperabilidade de SIs. O sucesso das

⁹ KOS é o acrônimo inglês para *knowledge organization systems*.

¹⁰ Vide iniciativa *OBO Foundry*. Disponível na Internet em: <http://www.obofoundry.org/>. Acesso em: 29/07/2013

ontologias na área médica impulsionou o seu uso em outros campos como tentativa de resolver problemas de interoperabilidade no âmbito público e privado.

Entretanto, nem sempre está claro como as ontologias podem ser usadas para promover interoperabilidade de SIs. As seções seguintes buscam esclarecer essa questão, explicando quais as possibilidades de uso de ontologias para integrar SIs (seção 4.2) e apresentando um estudo de caso de uso de ontologias em modelagem (seção 4.3).

4.2) Que interoperabilidade as ontologias podem proporcionar?

O problema de interoperabilidade que pode ser tratado a partir de ontologias é a heterogeneidade semântica. Conforme mencionado anteriormente (seção 2), mesmo que SIs adotem a mesma sintaxe, ou seja, os mesmos termos para se referir a coisas do mundo, eles normalmente associam diferentes significados a esses termos, ou seja, usam semânticas diferentes. Esse fato impede a troca de direta de informação entre SIs. Para solucionar esse problema, é preciso uma forma de especificar, sem ambiguidade, os vocabulários subjacentes aos SIs. Ontologias são capazes de proporcionar tal especificação. O restante dessa seção explica como isso pode ser feito. Em primeiro lugar revisita-se, de forma breve, a questão da heterogeneidade semântica, para em seguida explicar o uso das ontologias como alternativa para solucionar o problema.

Para explicar o processo de a comunicação ou troca de informação, seja entre pessoas ou sistemas, utilizam-se aqui aspectos já amplamente conhecidos da teoria da informação de Shanon e Weaver (1949).

A comunicação ocorre a partir da troca de informação entre agentes, sendo um agente envia e o outro recebe. A informação é comunicada em uma língua. Qualquer linguagem consiste de um conjunto de símbolos organizados. Sozinhos, esses símbolos não tem significado. É preciso que os agentes envolvidos expliquem como eles devem ser interpretados. As linguagens nesse processo podem ser tanto linguagens naturais, usadas para comunicação entre pessoas; quanto linguagens formais usadas para comunicação entre computadores. Independentemente do tipo, uma linguagem é caracterizada por uma sintaxe e por uma semântica. A sintaxe corresponde aos símbolos da linguagem mais as regras para arranjar esses símbolos em sentenças bem formadas. A semântica estabelece o significado dos símbolos, ao especificar a interpretação desses símbolos na linguagem e em um contexto.

Uma classe de problemas na troca de informação ocorre porque o remetente e o receptor, por motivos diversos, usam diferentes semânticas. No âmbito dos SIs, o resultado é que o mesmo símbolo pode ter diferentes significados em diferentes linguagens. Essa questão básica, conhecida como polissemia, é assunto de pesquisa em Ciência da Informação e Linguística e inclui, além do caso do mesmo termo com diferentes significados, problemas causados por símbolos diferentes com o mesmo significado. Existem ainda casos em que o significado dos símbolos não são iguais, mas também não são diferentes: eles se sobrepõem em alguma medida.

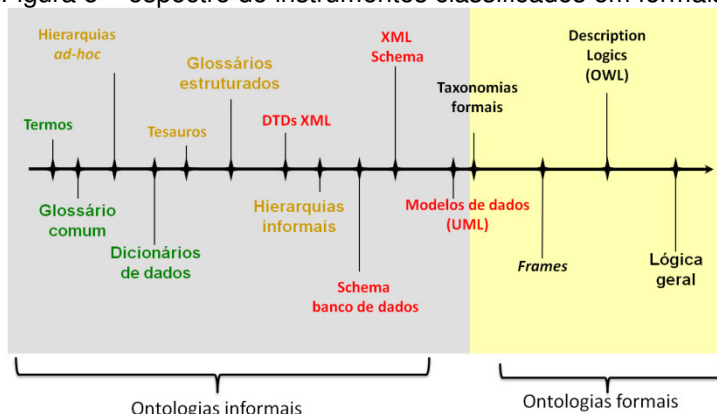
Quando a interação ocorre entre dois agentes humanos a polissemia é resolvida por interação recursiva, em linguagem natural, até que a questão seja resolvida. Mas isso não é possível quando os agentes são sistemas computacionais que atuam sem intervenção humana. Nesse caso, não existe uma solução simples, que resolva a heterogeneidade semântica. Não existe forma de interpretar os símbolos ao longo da comunicação, pois sistemas computacionais não têm capacidade de interpretação.

Nesse contexto, uma possibilidade para integrar dois diferentes vocabulários, V1 e V2, associados a dois SIs que atuam em domínios diferentes é estabelecer relações semânticas entre os termos de V1 e os termos de V2. Para fazer isso, é preciso definir o significado de cada termo de V1 e de V2 em uma linguagem que seja mais expressiva que os próprios V1 e V2. Essa linguagem deve expressar explicitamente o significado dos termos e evitar as ambiguidades inerentes à linguagem natural. Nesse ponto, ontologias podem ser usadas com vantagens.

As ontologias são instrumentos aptos a especificar explicitamente a semântica de termos pertencentes a diferentes domínios. Por isso podem proporcionar a troca de informação entre sistemas, e até mesmo entre pessoas (JASPER e USCHOLD, 1999). As ontologias porém podem ser construídas em diferentes níveis de formalidade. A Figura 5 apresenta diversos tipos de instrumentos, classificados aqui como tipos de ontologias, formais ou informais. Essa classificação não é consensual, mas objetiva aqui apenas destacar o nível de formalidade dos instrumentos.

Cabe então analisar como ontologias formais e ontologias informais podem atuar na busca por interoperabilidade, uma vez que o nível de formalidade vai impactar em sua capacidade de solucionar problemas de comunicação.

Figura 5 – espectro de instrumentos classificados em formais e não formais



Fonte: adaptado de Almeida, Souza e Fonseca (2011)

Quando a semântica dos vocabulários subjacentes aos SIs é especificada por uma *ontologia formal*, os termos e as relações entre os termos do vocabulário aparecem como variáveis, constantes e predicados de uma linguagem lógica formal. Axiomas lógicos são então adicionados a ontologia, de forma a expressar relações entre entidades ali representadas. Os axiomas são meios de restringir o significado dos termos: aceitam-se certas interpretações e rejeitam-se outras. Dessa forma a semântica de cada termo é definida explicitamente e sem ambiguidades (GUARINO, 1998).

A ontologia define o significado de termos e de relações entre termos ao explicitar como devem ser interpretadas. Os axiomas da ontologia podem ser usados, por exemplo, para explicar o sentido pretendido para a relação *parte-todo*, uma importante relação semântica, estabelecendo suas propriedades. Uma propriedade importante da relação *parte-todo* para fins de recuperação da informação é a transitividade, a qual se expressa da seguinte forma: se A é parte de B e B é parte de C, então A é parte de C. Entretanto, a transitividade das relações *parte-todo* não é totalmente consensual. Existem casos em que a transitividade é clara, como na cadeia transitiva apresentada por Winston et al. (1987): *O dedo de Simpson é parte da mão de Simpson; A mão de Simpson é parte do corpo de Simpson; O dedo de Simpson é parte do corpo de Simpson*. Mas existem casos em que a transitividade não parece razoável, como no exemplo fornecido por Lyons (1977): *A maçaneta é parte da porta; A porta é parte da casa; A maçaneta é parte da casa*. Os axiomas da ontologia formal podem, nesse caso, especificar o significado da relação, sua semântica, rejeitando interpretações que estão em desacordo com o significado pretendido para termos e relações do vocabulário. Assim, os agentes nas duas pontas do processo de comunicação vão lidar com a mesma semântica.

Ontologias informais também são usadas para especificar a semântica dos vocabulários subjacentes aos SIs. Exemplos são linguagens semi-formais, como a UML (OMG, 2012), usadas na modelagem de sistemas; e vocabulários controlados como os tesouros (ANSI, 2005). Instâncias de ontologias informais são os instrumentos já citados antes (seção 3): os padrões médicos, por exemplo, o HL7 e o SNOMED; e os vocabulários criados no âmbito do governo brasileiro, por exemplo, o VCGE. Esses instrumentos em geral tentam especificar o significado dos termos e relações ao fazer corresponder todas as possibilidades de definição à um conjunto de definições pré-estabelecidas em um domínio de interpretação limitado, fixo e consensual. A eliminação da ambiguidade dos termos é possível em alguma medida porque casos em que o mesmo símbolo tem diferentes significados são eliminados por acordo anterior sobre como usar os termos. Porém, as ontologias informais não possuem axiomas e assim, não tem como restringir o significado dos termos de forma a rejeitar ou aceitar certas interpretações, em um ambiente computacional.

Por exemplo, no caso do VCGE, definições para os termos, mesmo que informais, não estão disponíveis. Para entender o significado de um termo, é preciso se remeter à organização taxonômica do vocabulário. É possível então formar definições absurdas: “monitoramento de florestas”, bem como “exploração de floresta” e reposição de florestas”, os quais são claramente *tipos de processos*, são definidos como um tipo de “floresta”, o qual é claramente um *tipo de objeto*. Além da confusão ontológica, esse tipo de situação viola claramente qualquer proposta razoável de classificação. A “informalidade” na definição de termos que vão ser usados em vocabulários de SIs impossibilita a interpretação clara do significado pretendido e mina qualquer tentativa de interoperabilidade automática.

Como ontologias informais podem contribuir para a interoperabilidade? De forma limitada, e apenas na medida em que, geralmente, tem seu uso regulamentado pelo governo passando a ser obrigatórios por força de lei. A especificação da semântica de vocabulários por meio de ontologias informais pode ser suficiente para comunicação humana, mas computadores não tem como lidar com contexto e com conhecimento de senso comum. Por essa razão, para que a ontologia atue na comunicação entre máquinas é necessário especificar uma linguagem lógica formal que suporte inferências.

Finalmente, cabe sumarizar as possibilidades de interoperabilidade com o uso de ontologias em duas soluções principais, de forma a prover entendimento sobre seu

alcance. Essas duas possibilidades respondem a pergunta que deu título a esta seção: que interoperabilidade as ontologias podem proporcionar?

Um primeiro tipo de solução adota ontologias informais. Nesse caso, todas as aplicações envolvidas no processo de comunicação compartilham uma terminologia comum. A semântica dessa terminologia compartilhada é em geral especificada por um padrão, ou melhor, um meta-padrão. Todas as aplicações que aderem a esse meta-padrão podem se comunicar usando a mesma terminologia, sem problemas com ambiguidades, num modelo que é normalmente denominado federativo (ARMS, 2000). No caso de uma aplicação X que usa outra terminologia, diferente da terminologia especificada no meta-padrão, será necessário estabelecer mapeamentos entre as duas terminologias. Se a terminologia da aplicação X tem uma semântica bem definida, a questão da heterogeneidade semântica poderá ser resolvida por uma equipe de especialistas e programadores que criarão mapeamentos entre as duas terminologias. Essa é uma solução conhecida e resulta em esforço de programação para construir uma nova camada de *middleware* (LINTHICUM, 2002).

Em um segundo tipo de solução, as aplicações usam diferentes terminologias cuja semântica é especificada por ontologias formais. Uma terminologia genérica, cuja semântica também é especificada por uma ontologia formal, é usada como referência, como uma língua franca. Os mapeamentos entre as terminologias específicas dos sistemas envolvidos na comunicação, nesse caso, são indiretos: cada terminologia é mapeada para a terminologia de referência, ou a partir dela. Uma vez que a semântica das terminologias específicas é também especificada por ontologias formais, os mapeamentos podem então ser computados automaticamente (BECHHOFFER e GOBLE 2001). Essa abordagem que permite a geração de mapeamentos pelos sistemas envolvidos no processo de comunicação, ou seja, que permite que a tradução entre as semânticas das terminologias envolvidas seja feita automaticamente, é essencialmente a proposta da web semântica como concebida foi por Berners-lee et al. (2001).

4.3) Aplicações reais de ontologias na busca por interoperabilidade

Exemplos práticos de uso de ontologias em busca de interoperabilidade têm sido conduzidos como pesquisa já há algum tempo na Universidade Federal de Minas Gerais (OLIVEIRA e ALMEIDA, 2011) e recentemente em conjunto com a Companhia de

Tecnologia da Informação do Estado de Minas Gerais (Prodemge). O restante da presente seção apresenta caso de estudo em andamento.

O contexto de desenvolvimento da pesquisa abrange modelos UML reais de projeto do Governo do Estado de Minas Gerais, o GRP Minas (*Government Resource Planning*), especificamente o processo de negócio denominado Plano Plurianual de Ação Governamental (PPAG), um instrumento normatizador do planejamento público.

Os instrumentos utilizados para condução da pesquisa são a ontologia *Unified Foundational Ontology* (UFO), a linguagem de modelagem OntoUML e a ferramenta OLED. Na verdade, esses três instrumentos foram criados conjuntamente como um arcabouço teórico-prático para modelagem (GUIZZARDI et al. 2009).

A UFO é uma ontologia de referência desenvolvida especificamente para dar suporte à modelagem. Reúne teorias axiomáticas que abordam as principais categorias de conceitos usados em modelagem conceitual. No âmbito do arcabouço UFO, foi criada uma linguagem de modelagem conceitual, a qual adota como primitivas de modelagem as restrições ontológicas propostas por axiomas da UFO. Essa linguagem, denominada OntoUML, foi especificada acima do metamodelo da UML 2.0 (OMG, 2012). O objetivo era garantir o mapeamento entre a UML e a estrutura definida pela UFO. Além disso, foi criada uma ferramenta, um editor para criação de modelos na linguagem OntoUML, denominado *OntoUML Lightweight Editor* (OLED)¹¹.

A pesquisa avalia modelos durante o desenvolvimento do sistema, de forma a identificar possíveis inconsistências de modelagem. As etapas de avaliação são: i) mapeamento do modelo do PPAG para a OntoUML; ii) verificação da aderência do modelo ao padrão ontológico da UFO por meio da ferramenta OLED; e iii) correção do modelo do PPAG e nova verificação no OLED, em um processo recursivo, até a eliminação de erros.

Durante o mapeamento, uma característica comumente observada nos modelos de domínio do projeto foi a ampla utilização de relações parte-todo. Dessa forma, as primeiras iniciativas se voltaram para avaliação de possíveis inconsistências na criação desse tipo de relação. O modelo foi então mapeado para a OntoUML através da ferramenta OLED.

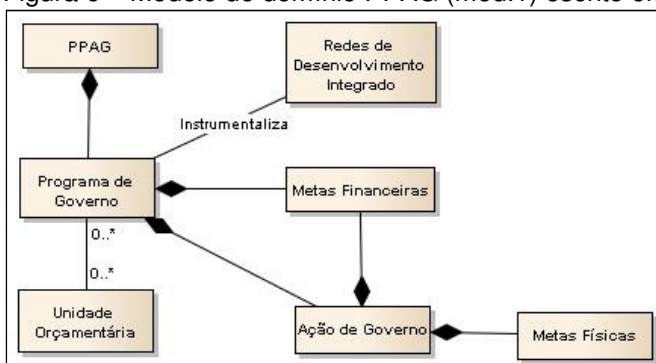
Nos casos em que a ferramenta acusava erros de modelagem nos resultados da verificação, o modelo em questão era reconstruído de acordo com a documentação dos

¹¹ Disponível na Internet em: <<http://code.google.com/p/ontouml-lightweight-editor/>>. Acesso em 20 de agosto de 2012.

seus requisitos e com o conhecimento de especialistas do projeto. Em seguida o modelo era novamente testado, até que a ferramenta não apontasse mais erros de modelagem.

Na Figura 6 é possível observar o modelo que a equipe de analistas do projeto desenvolveu, chamado aqui mod.1. Na sequência, o mod.1 foi reescrito em linguagem OntoUML gerando o mod.2, o qual foi finalmente importado pela ferramenta OLED e submetido à validação.

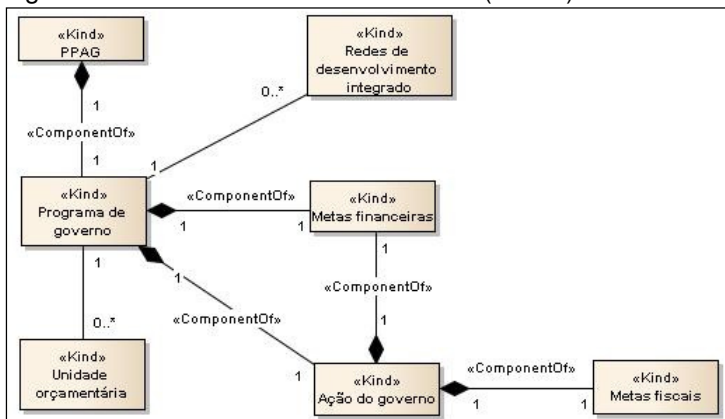
Figura 6 – Modelo de domínio PPAG (mod.1) escrito em UML



Fonte: Projeto PPAG

Em seguida, foi realizada uma análise dos modelos mod.1 e mod.2. Observaram-se decisões que não atendiam a restrições ontológicas, revelando problemas de modelagem recorrentes no modelo original. Por exemplo, uma *ação não é parte de um programa*, assim como *um programa não é parte do PPAG*.

Figura 7 – Modelo de domínio do PPAG (mod.2) reescrito em OntoUML



Fonte: Projeto PPAG

A ferramenta OLED apontou problemas de validação conforme apresentado na Figura 8 e o modelo foi então reescrito utilizando a OntoUML (mod.3, Figura 9).

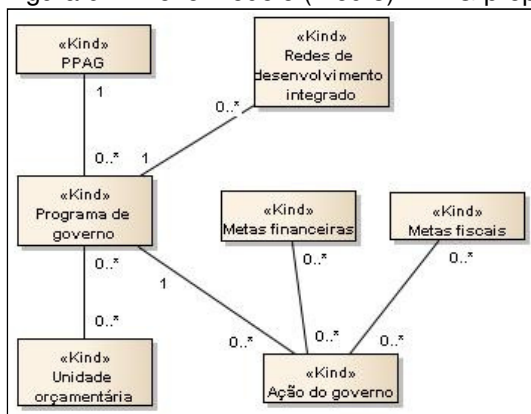
Figura 8 – Resultado de validação do modelo PPAG pelo OLED

ERROR: The model is not valid sintatically. The following error(s) where found:
 [unnamed] («componentof») - The sum of the minimum cardinalities of the parts must be greater or equal to 2
 Model validated in 128 ms, 1 error(s) found

Fonte: saída da ferramenta OLED

As relações que foram identificadas no modelo de domínio constituíam-se na verdade em *relações de associação*, não em *relações parte-de*. Assim, foi proposto um novo modelo de domínio para o assunto em questão, denominado aqui mod.3 (Figura 9), o qual foi também submetido ao OLED e então considerado válido.

Figura 9 – Novo modelo (mod.3) PPAG proposto em OntoUML



Fonte: saída da ferramenta OLED.

5) Considerações Finais

O presente artigo definiu tipos de interoperabilidade e descreveu brevemente as iniciativas brasileiras para solução para dificuldades de integração entre sistemas. Descreveu a aplicação de ontologias como alternativa para lidar com falta de interoperabilidade, distinguindo duas possibilidades: o uso de ontologias informais e o uso de ontologias formais. Finalmente, apresentou um caso real de pesquisa em andamento, no qual modelos de sistemas são incrementados via a axiomatização de uma ontologia formal. Cabe agora apresentar algumas considerações finais e conclusões.

Uma primeira consideração necessária é explicar que tipo de solução é aquela apresentada como estudo de caso (seção 4.3). Onde está solução se localiza em relação as duas alternativas apresentadas para uso de ontologias? Trata-se de uma solução intermediária, que faz uso de princípios de uma ontologia formal para reduzir inconsistências em modelos de Sis. Mas os mapeamentos não são automáticos porque uma das pontas do processo – o modelo UML sob avaliação – não é uma ontologia formal. Ainda assim, alguma automatização foi obtida pela ferramenta que testa o modelo a partir de princípios da ontologia formal e apresenta mensagens de erro.

Uma segunda consideração importante diz respeito exatamente a inferências. As inferências são a chave para entender a diferença entre as propostas de uso de ontologias formais e informais, bem como em que medida cada solução pode ser útil com respeito

a interoperabilidade. Por limitações de espaço, não se explicou muito sobre inferências aqui, mas para uma visão geral sugere-se Antoniou e Van Harmelen (2004).

As inferências, características das ontologias formais, permitem que o processo de classificação seja incrementado a partir da criação de relações taxonômicas de forma automática. Esse tipo de abordagem não é exatamente uma novidade (BECHHOFFER e GOBLE 2001), mas sua aplicação torna-se cada vez mais atrativa a medida que aumenta exponencialmente o volume dos dados heterogêneos que precisam ser manipulados por SIs. A possibilidade de que mapeamentos necessários para a interoperabilidade sejam computados automaticamente é parte da pesquisa em Representação do Conhecimento. Entretanto, não está claro ainda a relação custo-benefício entre o esforço requerido para formalização e os resultados obtidos em interoperabilidade. Independentemente disso, é essa a proposta que as ontologias oferecem em termos de interoperabilidade.

Conclui-se que ontologias podem desempenhar papel relevante para obtenção da interoperabilidade entre sistemas, mas avaliando-se as iniciativas nacionais, muito ainda está por ser feito para que essa possibilidade se torne real. Isso ocorre porque essas iniciativas enfatizam ainda a criação ou o estabelecimento de padrões técnicos.

Um primeiro movimento no sentido de atacar outras frentes é o tipo de proposta apresentada como estudo de caso (seção 4.3), onde se busca melhorar práticas de modelagem com o uso de ontologias formais. Esse tipo de providência, por si só, já traria grande benefícios a iniciativas nacionais como o vocabulário VCGE, o qual padece da não adoção de princípios consolidados de organização do conhecimento que a Ciência da Informação oferece.

Um segundo movimento pode remeter ao fato de que o uso de ontologias formais parece ser um sucesso na área de medicina ao redor do mundo. Assim, as iniciativas nacionais talvez possam privilegiar essa área tão importante para a sociedade. De imediato, é preciso também avaliar criticamente soluções adotadas no Brasil, como por exemplo o padrão de registros médicos já mencionado (seção 3), o openEHR. Para uma avaliação crítica desse padrão do ponto de vista ontológico e com vistas a interoperabilidade sugere-se consultar Andrade, Almeida e Schulz (2012).

Espera-se que o artigo tenha lançado alguma luz sobre a importância da questão da interoperabilidade, sobre o que tem sido feito, sobre o que é necessário fazer e como a Ciência da Informação pode contribuir. Em pesquisas futuras, espera-se que dados empíricos possam comprovar a utilidade das soluções apresentadas aqui.

Referências

- ALMEIDA, M. B. Revisiting Ontologies: a necessary clarification. *Journal of the American Society of Information Science and Technology*. v. 64, n. 8. p. 1682-93. 2013.
- ALMEIDA, M.; SOUZA, R.; FONSECA, F. (2011). Semantics in the semantic web: A critical evaluation. *Knowledge Organization*, v. 38, n. 3, p. 187–203, 2011.
- ANSI/NISO Z39.19-2005. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. Bethesda: National Information Standards Organization Press, 2005.
- ANDRADE, A.Q.; ALMEIDA, M.B.; SCHULZ, S. *Revisiting ontological foundations of the OpenEHR Entry Model*, 2013. Disponível na Internet em: <<http://mba.eci.ufmg.br/pu.html>>. Acesso em: 22 de maio de 2013.
- ANTONIOU, G.; Van HARMELEN, F. 2004. *A Semantic Web Primer*. Massachusetts: The MIT Press, 2004.
- ARMS, W. Y. *Thoughts about interoperability in the NSDL: draft for discussions*, 2000. Disponível em: <<http://www.cs.cornell.edu/wya/papers/NSDL-Interop.doc>>. Acesso em: 05 de junho de 2013.
- ARMS, W.Y. et al. *A spectrum of interoperability: the site for science for prototype for the NSDL*, 2002. Disponível em: <<http://www.dlib.org/dlib/january02/arms/01arms.html>>. Acesso em: 05/06/2013.
- BECHHOFFER, S.; GOBLE, C. Thesaurus construction through knowledge representation. *Data & Knowledge Engineering*, v.37, p. 25-45, 2001.
- BERNERS-LEE et al. *The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, 2001. Disponível em: <<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>>. Acesso em 10 março 2010.
- BRASIL, Ministério da Saúde Gabinete do Ministro. *Portaria nº 2.073, de 31 de agosto de 2011*. Disponível na Internet em: <<http://www.brasilsus.com.br/legislacoes/gm/109456-2073.html>>. Acesso em 30 de abril de 2013.
- BISHR, Y. *Semantic aspect of interoperable GIS*, 1997. Disponível na Internet em: <<http://library.wur.nl/WebQuery/wda/947563>>. Acesso em 20 de julho de 2009.
- ePING. *Padrões de Interoperabilidade de Governo Eletrônico*. Documento de Referência da e-PING Versão 2013, 2013. Disponível na Internet em <<http://www.governoeletronico.gov.br/biblioteca/arquivos/documento-da-e-ping-versao-2013/>>. Acesso em 8 de julho de 2013.
- ePMG. *Padrão de Metadados do Governo Eletrônico*, 2010. Disponível na Internet em: <<http://www.governoeletronico.gov.br/biblioteca/arquivos/e-pmg>>. Acesso em 20 de julho de 2013.

- GPG – *Guia de Gestão de Processos de Governo*, 2011. Disponível na Internet em: <<https://www.governoeletronico.gov.br/biblioteca/arquivos/guia-de-gestao-de-processos-de-governo/download>>. Acesso em: 20 de maio de 2013.
- GRUBER, T. *What is an ontology?* 1993. Disponível na Internet em: <<http://www.wksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: 20 de agosto de 2009.
- GUARINO, N. Formal Ontology in Information Systems. *Proceedings of FOIS'98*, 1998. Amsterdam: IOS Press
- JASPER, R.; USCHOLD, M. *A framework for understanding and classifying ontology applications*, 1999. Disponível na Internet em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.6456>>. Acesso em: 20 de novembro de 2010.
- LANDSBERGEN, D.; WOLKEN, G. *Realizing the Promise: Government Information Systems and the Fourth*, 2001. Disponível na Internet em <<http://onlinelibrary.wiley.com/doi/10.1111/0033-3352.00023/abstract>>. Acesso em 22 de julho de 2010.
- GUIZZARDI, G. et al. *Ontologias de Fundamentação e Modelagem Conceitual*, 2009. Disponível em: <<http://nemo.inf.ufes.br/>>, Acesso em 10 de setembro de 2010.
- LINTHICUM, D. S. *Enterprise Application Integration*. New York: Addison-Wesley Professional, 2002.
- LYONS, J. *Semantics*: 1. Cambridge: Cambridge Press, 1977.
- MILLER, P. *Interoperability*. What is it and why should I want it? 2000. Disponível em: <<http://www.ariadne.ac.uk/issue24/interoperability/>>. Acesso em: 05/06/2013.
- MINISTÉRIO DA SAÚDE. *Portaria nº 2.073, de 31 de agosto de 2011*. Disponível em <http://bvsms.saude.gov.br/bvs/saudelegis/gm/2011/prt2073_31_08_2011.html>. Acesso em 20 dez. 2011.
- MOREIRA, W.; LARA, M. Ontologias, categorias e interoperabilidade semântica. *DataGramaZero, Revista de Informação*, v.13, n. 4, 2012.
- OLIVEIRA, V. N. P.; ALMEIDA, M. B. Um roteiro para avaliação ontológica de modelos de sistemas de informação. *Perspectivas em Ciência da Informação*, v. 16, n.1, p.165-184, 2011.
- OMG (Object Management Group). *OMG's MetaObject Facility*, 2012. Disponível na Internet em: <<http://www.omg.org/mof/>>. Acesso em: 23 janeiro de 2013.
- SANTOS, E. M. *Desenvolvimento e implementação da arquitetura e-PING: estratégias adotadas e possíveis implicações*, 2010. Disponível na Internet em: <www.teses.usp.br/teses/.../TeseErnaniMSantos.pdf >. Acesso em: 19 de julho de 2013.
- SANTOS, K. C. P. *Utilização de ontologias de referência como abordagem para Interoperabilidade entre sistemas de informação utilizados ao longo do ciclo de vida de produtos*. Dissertação Mestrado. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Eng. Mecânica e de Materiais, Curitiba, 2011. 181 p.
- SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. *TransInformação*, v. 20, n. 2, p.133-148, 2008.

- SHANNON, C. E.; WEAVER, W. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press, 1949.
- SHETH, A. Changing Focus on Interoperability in Information Systems: from System, Syntax, structure to Semantics, 1999. Disponível na Internet em: <<http://lsdis.cs.uga.edu/lib/download/S98-changing.pdf>>. Acesso em: 17 de abril de 2011.
- SIMON, J.; SMITH, B. *Using Philosophy to Improve the Coherence and Interoperability of Applications Ontologies: A Field Report on the Collaboration of IFOMIS and L&C*, 2004. Disponível na Internet em: <<http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-112/Simon.pdf>>. Acesso em: 23 de setembro de 2010.
- SOERGEL, D. *The rise of ontologies or the reinvention of classification*, 1999. Disponível na Internet em: <<http://www.dsoergel.com/cv/B70.pdf>>. Acesso em: 3 de março de 2008.
- SMITH, B. *Ontology and Information Systems*, 2004. Disponível em: <<http://www.ontology.buffalo.edu/ontology/>>. Acesso em: 10 de junho de 2013.
- UKOLN. *Interoperability focus: looking at interoperability*, 2005. Disponível na Internet em: <<http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>>. Acesso em: 04 de junho de 2013.
- VCGE – Vocabulário Controlado do Governo Eletrônico, 2011. Disponível na Internet em: <vocab.e.gov.br/2011/03/vcge>. Acesso em: 05 de junho de 2013.
- VICKERY, B.C. Ontologies. *Journal of Information Science*, v. 23. N. 4, p. 227-286, 1997.
- WINSTON, M. E. et al. A Taxonomy of Part-Whole Relations. *Cognitive Science*, v.11, n. 4, p. 417-444, 1987.